STRUCTURAL PROPERTIES OF SOLUTIONS OF TOTAL VARIATION REGULARIZATION PROBLEMS

WOLFGANG RING

Abstract. In dimension one it is proved that the solution of a total variation-regularized least-squares problem is always a function which is "constant almost everywhere", provided that the data are in a certain sense outside the range of the operator which is to be inverted. A similar, but weaker result is derived in dimension two.

1. Introduction

Since the work by Rudin, Osher and Fatemi [13] appeared in 1992, regularization by total variation functionals has received considerable attention in image and signal processing (cf. Chambolle and Lions [3], Acar and Vogel [1], Dobson and Santosa [6], Vogel and Oman [15], Chavent and Kunisch [4], Ito and Kunisch [10], Dobson and Scherzer [5], Nashed and Scherzer [12] and the references cited therein).

In this paper, we consider the least-squares problem

$$\min_{u \in BV(\Omega)} \|Ku - z\|^2 + \alpha |\nabla u|(\Omega), \tag{1.1}$$

where $|\nabla u|(\Omega)$ stands for the total variation of the (distributional) gradient of u. The variational problem (1.1) should be considered as a stable approximation to the, possibly ill-posed problem

$$Ku = z.$$

It has been observed by several authors that "blocky" structures (i.e. piecewise constant functions) can be reconstructed very well by total variation regularization (cf. [6]) and, on the other hand, solutions of total variation-regularized problems are usually rather "blocky", even if the exact solution u_0, corresponding to exact (i.e. noise-free) data z_0, is smooth (cf. [3], [11]).

We illustrate the situation in two numerical examples. Here our goal is not the best possible reconstruction of the ideal solution. We choose configurations of noise and regularization parameter which display stuctural properties of the solution as good as possible.

We first consider a one-dimensional problem of type (1.1). Here K is an integral operator of the first kind,

$$Ku(x) = \int_0^x u(\xi)\, d\xi - \int_0^1 (1 - \xi) u(\xi)\, d\xi.$$

Date: August 1999.

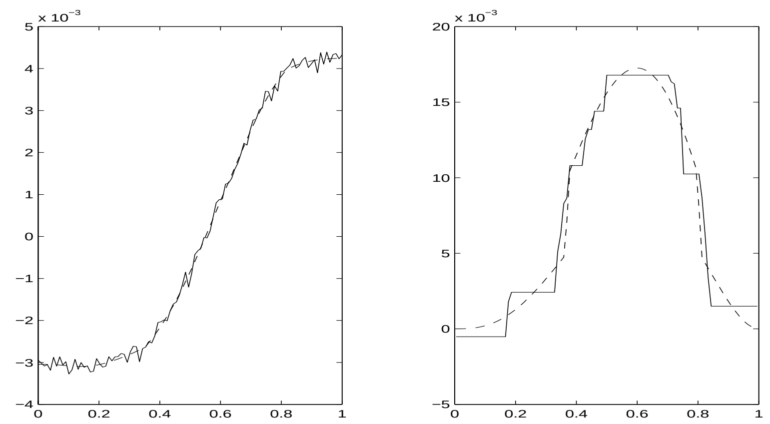

FIGURE 1. Unperturbed and perturbed data and solution to (1.1)

The plot on the left hand side in Figure 1 shows exact data z_0 together with noise-corrupted data z. We added artificial Gaussian noise to the data-vector z_0 to obtain z. The solution to (1.1) and the exact solution u_0 with $Ku_0 = z_0$ are plotted on the right hand side as solid and dashed lines respectively. It is apparent, that the solution $\overline{u}$ of (P) is constant over large subintervals and that there are jump discontinuities at a number of points.

Similar effects can be observed in dimension two. Figure 2 shows the regularized inversion of noise corrupted data $z \in L^2([0,1] \times [0,1])$ when K is the compact Fredholm operator

$$Kf(\mathbf{x}) = \int_\Omega e^{-2|\mathbf{x}-\boldsymbol{\xi}|^2} f(\boldsymbol{\xi})\, d\boldsymbol{\xi}.$$

with $\Omega = [0,1] \times [0,1]$. In this case, the solution is the union of disjoint patches where on each patch the gradient of the solution is pointing in the direction of one of the coordinate axes.

In this paper, give some theoretical explanations for this behavior. In dimension one we can characterize the structure of the solution in the presence of data-noise quite thoroughly. In this situation, we prove that the derivative of the solution is zero almost everywhere. In dimension two we do not obtain such a strong result, but we can derive some structural properties of solutions, which correspond to the above mentioned observation that, at each point, the gradient of the solution is parallel to one of the coordinate directions. In both, the one- and the two-dimensional situation, we focus our attention on two important special cases. The problem of denoising noise-corrupted data and the problem of deblurring data which are blurred by some smoothing filter.

In Section 2 we introduce the spaces $BV(\Omega)$ and $\mathcal{M}_n(\Omega)$ and we recall some structural properties of functions of bounded variation and of Radon measures, where we use mainly the Lebesgue decomposition of a measure in its absolutely continuous and its singular part. In Section 3, we formulate the

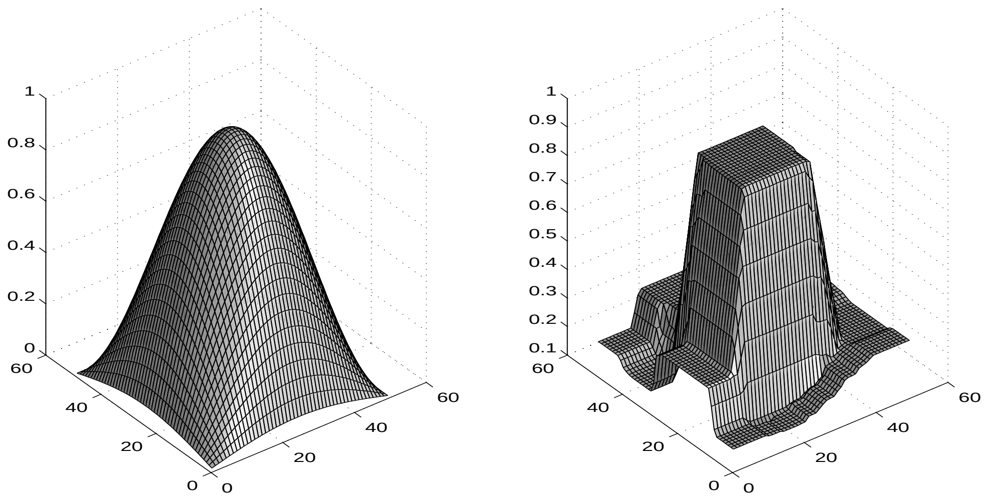

FIGURE 2. Ideal solution and solution to (P) for a two-dimensional Problem

variational problem and its functional analytic setting, we prove existence of a solution and we derive an optimality system for (1.1). In Sections 4 and 5 we study structural properties of solutions to (1.1) in dimensions one and two respectively.

2. THE SPACES $BV(\Omega)$ AND $\mathcal{M}_n(\Omega)$.

Let Ω be a bounded, open, simply connected domain in $\mathbb{R}^n$ with $n \geq 1$. We consider the space

$$BV(\Omega) = \{f \in L^1(\Omega) \, : \, \frac{\partial f}{\partial x_i} \in \mathcal{M}(\Omega) \text{ for } i = 1, \dots, n\}, \tag{2.1}$$

where $\frac{\partial}{\partial x_i}$ denotes the distributional derivative with respect to x_i and $\mathcal{M}(\Omega)$ is the vector space of all Radon measures on Ω, i.e. the dual of $\mathcal{C}_0(\Omega)$ (cf. Rudin, [14, Thm.6.19,p.130]). $BV(\Omega)$ is a Banach space if we define the norm on $BV(\Omega)$ by

$$\|f\|_{BV} = \|f\|_{L^1} + \sum_{i=1}^{n} \left| \frac{\partial f}{\partial x_i} \right| (\Omega), \tag{2.2}$$

where $\left| \frac{\partial f}{\partial x_i} \right| (\Omega)$ denotes the total variation of the measure $\frac{\partial f}{\partial x_i}$. We shall write (2.2) in the short form

$$\|f\|_{BV} = \|f\|_{L^1} + |\nabla f|(\Omega). \tag{2.3}$$

The term $|\nabla f|(\Omega)$ can alternatively be written as

$$|\nabla f|(\Omega) = \sup\Big\{ \int_\Omega f(\mathbf{x}) \operatorname{div} \mathbf{v}(\mathbf{x})\, d\mathbf{x} \,:\, \mathbf{v} \in \mathcal{C}_0^\infty(\Omega)^n,\ |\mathbf{v}(\mathbf{x})|_\infty \le 1 \text{ for all } x \in \Omega \Big\}. \tag{2.4}$$

Remark 1. If we choose a different vector norm for $\mathbf{v}$ in (2.4), we get a different (but equivalent) norm on $BV(\Omega)$. We choose the $|\cdot|_\infty$-norm, because then (2.2) and (2.4) are equivalent definitions. For any other norm for $\mathbf{v}$, the term corresponding to $|\nabla f|(\Omega)$ cannot be expressed as simple sum of variations, or any other simple function of $\left|\frac{\partial f}{\partial x_i}\right|(\Omega)$. Using our definition, we get a norm, which is easy to analyze, but, from the practical point of view, with the drawback of being anisotropic.

Remark 2. For $u \in BV(\Omega)$ with $\frac{\partial u}{\partial x_i} \in L^1(\Omega) \subset \mathcal{M}(\Omega)$, we have (Rudin [14, Thm.6.13,p.125])

$$\left|\frac{\partial u}{\partial x_i}\right|(\Omega) = \int_\Omega \left|\frac{\partial u}{\partial x_i}(\mathbf{x})\right| d\mathbf{x}.$$

Thus $W^{1,1}(\Omega)$ is isometrically (but not densely) embedded in $BV(\Omega)$.

We have the following properties.

Proposition 1.

1. Lower Semicontinuity

   If $\{u_j\}_{j=1}^\infty \subset BV(\Omega)$ *and* $u_j \to u$ *in* $L^1(\Omega)$ *then*

$$\left|\frac{\partial u}{\partial x_i}\right|(\Omega) \le \liminf_{j\to\infty} \left|\frac{\partial u_j}{\partial x_i}\right|(\Omega). \tag{2.5}$$

2. Compactness

   For every bounded sequence $\{u_j\}_{j=1}^\infty \subset BV(\Omega)$ *and for every* $p \in [1, \frac{n}{n-1})$ *there exists a subsequence* $\{u_{j_k}\}_{k=1}^\infty$ *and a function* $u \in BV(\Omega)$ *such that* $u_{j_k} \to u$ *in* $L^p(\Omega)$. *If* $n = 1$, *"*$\frac{n}{n-1}$*" has to be replaced by "*∞*".*

3. Sobolev Inequality

   There exists a constant $C = C(n, \Omega)$ *such that*

$$\|u - \tilde{u}\|_{L^{\frac{n}{n-1}}(\Omega)} \le C\, |\nabla u|(\Omega), \text{ for all } u \in BV(\Omega), \tag{2.6}$$

   with $\tilde{u} = \frac{1}{meas(\Omega)} \int_\Omega u\, d\mathbf{x}$. *If* $n = 1$, *the* $L^{\frac{n}{n-1}}$*-norm is understood to be the* L^∞*-norm.*

For proofs of the above described properties we refer to Giusti [9], Chavent and Kunisch [4] and Ziemer [17].

We consider the space $\mathcal{M}_n(\Omega) = \overset{n}{\underset{i=1}{\otimes}} \mathcal{M}(\Omega)$. A norm on $\mathcal{M}_n(\Omega)$ is given by

$$\|\boldsymbol{\mu}\|_{\mathcal{M}_n} = \sum_{i=1}^{n} |\mu_i|(\Omega), \tag{2.7}$$

where $|\mu_i|(\Omega)$ denotes the total variation of the measure μ_i and

$$\boldsymbol{\mu} = (\mu_1, \dots, \mu_n).$$

Remark 3. We can equivalently define the norm on $\mathcal{M}_n(\Omega)$ by

$$|\boldsymbol{\mu}| = \sup\Big\{ \sum_{i=1}^{n} \int_\Omega v_i(\mathbf{x})\, d\mu_i(\mathbf{x}) \,:\, \mathbf{v} \in \mathcal{C}_0^\infty(\Omega)^n,\ |\mathbf{v}(\mathbf{x})|_\infty \le 1 \text{ for all } x \in \Omega \Big\}. \tag{2.8}$$

For $u \in BV(\Omega)$ we have $\nabla u \in \mathcal{M}_n(\Omega)$, and definition (2.8) coincides with (2.4).

Let $\mu \in \mathcal{M}(\Omega)$. Then it is possible to write μ in a unique way as

$$\mu = \mu_a + \mu_s, \tag{2.9}$$

with μ_a being absolutely continuous with respect to the Lebesgue measure and μ_s being concentrated on a set of Lebesgue measure 0 (cf. Rudin [14, Thm.6.10,p.121]). The relation (2.10) is called the Lebesgue decomposition of μ. For $\boldsymbol{\mu} \in \mathcal{M}_n(\Omega)$, we have the Lebesgue decomposition

$$\boldsymbol{\mu} = \boldsymbol{\mu}_a + \boldsymbol{\mu}_s, \text{ with } \boldsymbol{\mu}_a = (\mu_{1,a}, \dots, \mu_{n,a});\ \ \boldsymbol{\mu}_s = (\mu_{1,s}, \dots, \mu_{n,s}). \tag{2.10}$$

If $\mu_{i,s}$ is concentrated on the zero-set $S_i \subset \Omega$, then $\boldsymbol{\mu}_s$ is concentrated on $S = \cup_{i=1}^n S_i$, which is also of Lebesgue measure 0. Thus $\boldsymbol{\mu}_s$ is a vector-valued singular measure. Moreover $\boldsymbol{\mu}_s$ and $\boldsymbol{\mu}_a$ are unique with this property.

Lemma 1. *We have*

$$|\boldsymbol{\mu}|(\Omega) = |\boldsymbol{\mu}_a|(\Omega) + |\boldsymbol{\mu}_s|(\Omega), \tag{2.11}$$

where $\boldsymbol{\mu} = \boldsymbol{\mu}_a + \boldsymbol{\mu}_s$ *is the Lebesgue decomposition of* $\boldsymbol{\mu} \in \mathcal{M}_n(\Omega)$ *into its absolutely continuous and its singular part, as given in* (2.10).

Proof. We first prove the result for dimension $n = 1$. Obviously we have $|\mu|(\Omega) \le |\mu_a|(\Omega) + |\mu_s|(\Omega)$ by the triangle inequality on $\mathcal{M}(\Omega)$.

To prove the converse inequality suppose we have given two partitions $\{E_i\}_{i=1}^\infty$ and $\{F_n\}_{n=1}^\infty$ of Ω, i.e. two countable collections of measurable sets with $E_i \cap E_j = \emptyset$ it $i \ne j$; $F_n \cap F_m = \emptyset$ if $n \ne m$ and $\cup_i E_i = \cup_n F_n = \Omega$. Since μ_a and μ_s are mutually singular, there exist two measurable sets A and B with $A \cap B = \emptyset$ and $A \cup B = \Omega$ such that $\mu_a(E) = \mu(A \cap E)$ and

$\mu_s(E) = \mu(B \cap E)$ for every measurable set E. We have

$$\sum_{i=1}^{\infty} |\mu_a(E_i)| + \sum_{n=1}^{\infty} |\mu_s(F_n)| = \sum_{i=1}^{\infty} |\mu(E_i \cap A)| + \sum_{n=1}^{\infty} |\mu(F_n \cap B)| \le |\mu|(\Omega)$$

since $\{E_i \cap A\}_{i=1}^{\infty} \cup \{F_n \cap B\}_{n=1}^{\infty}$ is a partition of Ω. If we take the supremum over all partitions $\{E_i\}_{i=1}^{\infty}$ and $\{F_n\}_{n=1}^{\infty}$, we get

$$|\mu_a|(\Omega) + |\mu_s|(\Omega) \le |\mu|(\Omega).$$

The result for dimension $n > 1$ follows immediately from the definition of $|\boldsymbol{\mu}|(\Omega)$ and from the one-dimensional case. □

Remark 4. We can isometrically identify absolute continuous measures with L^1-functions. Hence we can write (2.9) in the form $\mu = f + \mu_s$ with some $f \in L^1(\Omega)$. This yields the following decomposition of $\mathcal{M}(\Omega)$ in a direct sum:

$$\mathcal{M}(\Omega) = L^1(\Omega) \oplus \mathcal{P} \tag{2.12}$$

where

$$\mathcal{P} = \{\mu_s \in \mathcal{M}(\Omega) \,:\, \mu_s \text{ is singular w.r.t. the Lebesgue measure on } \Omega\}.$$

The decomposition for the vector-valued case reads as

$$\mathcal{M}_n(\Omega) = L^1_n(\Omega) \oplus \mathcal{P}^n, \tag{2.13}$$

with $L^1_n(\Omega) = \bigotimes_{i=1}^{n} L^1(\Omega)$ and $\mathcal{P}^n = \bigotimes_{i=1}^{n} \mathcal{P}$.

We consider the operator

$$\nabla : BV(\Omega) \to \mathcal{M}_n(\Omega).$$

Obviously, ∇ is a bounded linear operator on $BV(\Omega)$. We define the dual operator

$$-\mathfrak{Div} : \mathcal{M}_n(\Omega)^* \to BV(\Omega)^*; \quad -\mathfrak{Div} = \nabla^*.$$

Here and for the rest of the paper '$*$' denotes the dual for both, operators and spaces.

3. Problem Formulation, Solvability, and Optimality Systems

Let $\Omega \in \mathbb{R}^n$ be as in Section 2, $p \in [1, \frac{n}{n-1})$ if $n \ge 2$ and $p \in [1, \infty)$ if $n = 1$. Suppose Y is a reflexive Banach space and $K : L^p(\Omega) \to Y$ is a bounded linear operator. We consider the following unconstrained optimization problem.

$$\text{minimize } \frac{1}{2}\|Ku - z\|_Y^2 + \alpha|\nabla u|(\Omega) \text{ over } u \in BV(\Omega). \tag{P}$$

With $p < \frac{n}{n-1}$, and Ω bounded, we have $BV(\Omega) \subset L^p(\Omega)$ (Proposition 1.3), hence the minimization over $BV(\Omega)$ in (P) makes sense. Moreover, the embedding $BV(\Omega) \hookrightarrow L^p(\Omega)$ is continuous.

We have the following solvability result.

Proposition 2. *Suppose K and Y are given as above with $\{const\} \notin \ker K$. Then, for every $z \in Y$ there exists a solution $\overline{u}$ to problem* (P)*. If K is injective and the norm on Y is strictly convex, then the solution is unique.*

Proof. Suppose $\{u_j\}_{j=1}^{\infty}$ is a minimizing sequence for problem (P). Since $\{Ku_j\}$ is bounded and Y is reflexive, it follows from the Eberlein-Shmulyan Theorem (Yosida, [16, p.141]), that there exists a weakly convergent subsequence (again denoted by the same symbol $\{u_j\}$) satisfying $Ku_j \rightharpoonup \overline{y}$, where '$\rightharpoonup$' denotes weak convergence in Y.

We prove that $\{u_j\}$ is bounded in $L^p(\Omega)$. From the Sobolev Inequality (2.6) we conclude that $\{v_j\} = \{u_j - \tilde{u}_j\}$ with $\tilde{u}_j = \frac{1}{\text{meas}(\Omega)} \int_\Omega u \, d\mathbf{x}$ is bounded in $L^{\frac{n-1}{n}}(\Omega)$ and hence also in $L^p(\Omega)$. We claim that also $\{\tilde{u}_j\}$ is bounded. Suppose otherwise that we have $\tilde{u}_j = r_j \chi_\Omega$ with $|r_j| \to \infty$ as $j \to \infty$. Then $\|K\tilde{u}_j\|_Y = |r_j| \, \|K\chi_\Omega\|_Y \to \infty$ as $j \to \infty$ since const $\notin \ker K$. Then, $\|Ku_j\|_Y \geq \|K\tilde{u}_j\|_Y - \|Kv_j\|_y$ implies that $\{Ku_j\}$ must be unbounded, which cannot hold true. Thus $\{u_j\} = \{v_j + \tilde{u}_j\}$ is bounded in $L^p(\Omega)$.

Since $\{|\nabla u_j|(\Omega)\}$ is bounded we conclude that $\{u_j\}$ is bounded in $BV(\Omega)$ and hence, by Proposition 1.2, there exists a subsequence $\{u_j\}$ satisfying $u_j \to \overline{u}$ in $L^p(\Omega)$ with $\overline{u} \in BV(\Omega)$. Note that we also have $u_j \to \overline{u}$ in $L^1(\Omega)$ by the continuous embedding $L^p(\Omega) \hookrightarrow L^1(\Omega)$. Since K is weakly closed, it follows that $\overline{y} = K\overline{u}$. Using the weak lower semicontinuity of the norm $\|\cdot\|_Y$ and the semicontinuity property of the BV-seminorm (Proposition 1.1) we find

$$\begin{aligned} \frac{1}{2}\|K\overline{u} - z\|_Y^2 + \alpha|\nabla \overline{u}|(\Omega) &\leq \liminf_{j\to\infty} \left(\frac{1}{2}\|Ku_j - z\|_Y^2 + \alpha|\nabla u_j|(\Omega) \right) \\ &= \inf_{u \in BV(\Omega)} \left(\frac{1}{2}\|Ku - z\|_Y^2 + \alpha|\nabla u|(\Omega) \right), \end{aligned}$$

which proves that $\overline{u}$ is a solution to (P).

To prove uniqueness of the minimizer, it is sufficient to show that the cost functional (P) is strictly convex. This, however, is an immediate consequence of the injectivity of K and the strict convexity of $\|\cdot\|_Y^2$.

□

We can characterize the solution $\overline{u}$ of (P) by the following optimality system.

Proposition 3. *Let the assumptions of Proposition 2 hold and assume further that Y^* is a strictly convex Banach space. Then $\overline{u}$ is a solution to* (P) *if and only if there exists $\overline{\lambda} \in \mathcal{M}_n(\Omega)^*$ satisfying*

$$(\mathrm{O}_1) \qquad \Big(K^* J(K\overline{u} - z), u\Big)_{L^q, L^p} - \alpha \langle \mathfrak{Div}\, \overline{\lambda}, u \rangle_{BV^*, BV} = 0$$

and

$$\langle \overline{\lambda}, \boldsymbol{\mu} - \nabla \overline{u} \rangle_{\mathcal{M}_n^*, \mathcal{M}_n} + |\nabla \overline{u}|(\Omega) \leq |\boldsymbol{\mu}|(\Omega) \tag{O$_2$}$$

for all $u \in BV(\Omega)$ *and* $\boldsymbol{\mu} \in \mathcal{M}_n(\Omega)$. *Here* $q \in [1, \infty)$ *is chosen such that* $\frac{1}{p} + \frac{1}{q} = 1$ *and* $J : Y \to Y^*$ *denotes the duality map on* Y, *i.e. the subdifferential of the convex function* $\frac{1}{2} \| \cdot \|_Y^2$ *on* Y *(cf. Barbu* [2, p.60] *for details).*

Proof. The cost functional

$$F(u) = \frac{1}{2} \|Ku - z\|_Y^2 + \alpha |\nabla u|(\Omega)$$

is a convex continuous function on $BV(\Omega)$, thus $\overline{u}$ is the solution to (P) if and only if $0 \in \partial F(\overline{u})$ where $\partial F(u)$ denotes the subdifferential of F at u (cf. Ekeland and Turnbull [7, Prop.1,p.124]). We have [7, Cor.1,p.121]

$$\partial F(u) = \partial F_1(u) + \alpha \, \partial F_2(u)$$

with $F_1(u) = \frac{1}{2} \|Ku - z\|_Y^2$ and $F_2(u) = |\nabla u|(\Omega)$. If follows ([7, Prop.12,p.119] and [2, Ex.1,p.60 and Thm.1.2,p.2]) that $\partial F_1(u)$ is single-valued due to the strict convexity of Y^* and

$$\langle \partial F_1(u), w \rangle_{BV^*, BV} = \Big(K^* J(Ku - z), w \Big)_{L^q, L^p}$$

for all $w \in BV(\Omega)$. It follows from the same reference [7, Prop.12,p.119] that $\psi \in \partial F_2(u)$ iff $\psi = -\mathfrak{Div}\, \lambda$ and

$$\lambda \in \partial \tilde{F}_2(\nabla u), \tag{3.1}$$

where $\tilde{F}_2 : \mathcal{M}_n(\Omega) \to \mathbb{R}$ is given by $\tilde{F}_2(\boldsymbol{\mu}) = |\boldsymbol{\mu}|(\Omega)$. If we take the characterizations of ∂F_1 and ∂F_2 together, we get (O$_1$). The relation (O$_2$) is just the definition of (3.1)

□

We now reformulate the optimality system (O), to get a more convenient set of equations, from which we can conclude structural properties of the solution $\overline{u}$. We start with the optimality condition (O$_2$). Setting $\boldsymbol{\mu} = 0$ and $\boldsymbol{\mu} = 2\nabla \overline{u}$, in (O$_2$), and using the fact that $\tilde{F}_2$ is positive homogeneous of degree 1, we get

$$\langle \overline{\lambda}, \nabla \overline{u} \rangle_{\mathcal{M}_n^*, \mathcal{M}_n} = |\nabla \overline{u}|(\Omega)$$

and hence, by inserting into (O$_2$),

$$\langle \overline{\lambda}, \boldsymbol{\mu} \rangle_{\mathcal{M}_n^*, \mathcal{M}_n} \leq |\boldsymbol{\mu}|(\Omega)|. \tag{3.2}$$

for all $\boldsymbol{\mu} \in \mathcal{M}_n(\Omega)$. We claim that equality holds in (3.2) not only for $\boldsymbol{\mu} = \nabla \overline{u}$, but also if we set $\boldsymbol{\mu} = (\nabla \overline{u})_a$ and $\boldsymbol{\mu} = (\nabla \overline{u})_s$, the absolute continuous and singular part of $\nabla \overline{u}$ in $\mathcal{M}_n(\Omega)$ respectively. Suppose otherwise that strict inequality holds in at least one relation. Then, using Lemma 1, we have

$$\begin{aligned}|\nabla\overline{u}|(\Omega) =& (|\nabla\overline{u}|)_a(\Omega) + (|\nabla\overline{u}|)_s(\Omega)\\ >& \langle\overline{\lambda}, (\nabla\overline{u})_a\rangle_{\mathcal{M}_n^*,\mathcal{M}_n} + \langle\overline{\lambda}, (\nabla\overline{u})_s\rangle_{\mathcal{M}_n^*,\mathcal{M}_n}\\ =& \langle\overline{\lambda}, \nabla\overline{u}\rangle_{\mathcal{M}_n^*,\mathcal{M}_n} = |\nabla\overline{u}|(\Omega).\end{aligned}$$

Thus we cannot have a strict inequality in any case. We therefore obtain

$$\langle\overline{\lambda}, (\nabla\overline{u})_a\rangle_{\mathcal{M}_n^*,\mathcal{M}_n} = |(\nabla\overline{u})_a|(\Omega) \tag{O_{2a}}$$

and

$$\langle\overline{\lambda}, (\nabla\overline{u})_s\rangle_{\mathcal{M}_n^*,\mathcal{M}_n} = |(\nabla\overline{u})_s|(\Omega). \tag{O_{2s}}$$

We now show, that we can replace the multiplier $\overline{\lambda} \in \mathcal{M}_n(\Omega)^*$ by a smoother function $\hat{\lambda}$. Here we closely follow Chavent and Kunisch [4, Thm.2.4]. We consider the restriction of the functional $\boldsymbol{\mu} \mapsto \langle\overline{\lambda}, \boldsymbol{\mu}\rangle_{\mathcal{M}_n^*,\mathcal{M}_n}$ to $L^1_n(\Omega) \subset \mathcal{M}_n(\Omega)$.

$$\hat{\lambda} = \overline{\lambda}|_{L^1_n(\Omega)}. \tag{3.3}$$

Since $\hat{\lambda}$ is a bounded linear functional on $L^1_n(\Omega)$, we have

$$\hat{\lambda} \in L^\infty_n(\Omega) \tag{3.4}$$

and, by (O_{2a}) and (3.2)

$$\langle\hat{\lambda}, (\nabla\overline{u})_a\rangle_{L^\infty_n,L^1_n} = \int_\Omega |(\nabla\overline{u})_a(\mathbf{x})|\, d\mathbf{x} \tag{3.5}$$

$$\langle\hat{\lambda}, \mathbf{f}\rangle_{L^\infty_n,L^1_n} \leq \int_\Omega |\mathbf{f}(\mathbf{x})|\, d\mathbf{x} \text{ for all } \mathbf{f} \in L^1_n(\Omega). \tag{3.6}$$

Recall that we have to use the norm $\|\mathbf{f}\|_{L^1_n} = \sum_{i=1}^n \int_\Omega |f_i(\mathbf{x})|\, d\mathbf{x}$ to have compatability with (O_{2a}) and (3.2). From (3.5) and (3.6), we conclude that

$$\hat{\lambda}_i(\mathbf{x}) \in \operatorname{sign}\left(\frac{\partial\overline{u}}{\partial x_i}\right)_a(\mathbf{x}) \quad \text{a. e. on } \Omega \text{ for } i = 1, \dots, n. \tag{3.7}$$

Recall the optimality condition (O_1). We have

$$\mathfrak{Div}\,\overline{\lambda} = \frac{1}{\alpha}K^*J(K\overline{u} - z) \in L^q(\Omega) \subset BV(\Omega)^*. \tag{3.8}$$

Since $\hat{\lambda} \in L^\infty_n(\Omega) \subset \mathcal{D}_n(\Omega)$, we can calculate the distributional divergence $\operatorname{div}\hat{\lambda}$. Suppose $\varphi \in \mathcal{D}(\Omega)$. Then we have

$$\begin{aligned}\langle\operatorname{div}\hat{\lambda}, \varphi\rangle_{\mathcal{D}',\mathcal{D}} = -\langle\hat{\lambda}, \nabla\varphi\rangle_{L^\infty_n,L^1_n} = -\langle\overline{\lambda}, \nabla\varphi\rangle_{\mathcal{M}_n^*,\mathcal{M}_n} =\\ \langle\mathfrak{Div}\,\overline{\lambda}, \varphi\rangle_{BV^*,BV} = \langle\mathfrak{Div}\,\overline{\lambda}, \varphi\rangle_{L^q,L^p}\end{aligned} \tag{3.9}$$

for all $\varphi \in \mathcal{D}(\Omega)$. Since $\mathcal{D}(\Omega)$ is dense in $L^q(\Omega)$, we find

$$\operatorname{div}\hat{\lambda} = \mathfrak{Div}\,\overline{\lambda} \in L^q(\Omega). \tag{3.10}$$

By inserting this into (O_1), we get the following necessary optimality condition.

Proposition 4. *Let the assumptions of Proposition 3 hold and let* $\overline{u} \in BV(\Omega)$ *be the solution to* (P). *Then there exists* $\hat{\lambda} \in L^\infty_n(\Omega)$*, with* $\operatorname{div}\hat{\lambda} \in L^q(\Omega)$*, such that*

$$(O_1') \qquad K^*J(K\overline{u} - z) = \operatorname{div}\hat{\lambda} \in L^q(\Omega)$$

$$(O_2') \quad \hat{\lambda}_i(\mathbf{x}) \in \operatorname{sign}\left(\frac{\partial \overline{u}}{\partial x_i}\right)_a(\mathbf{x}) \quad \textit{almost everywhere on } \Omega \textit{ for } i = 1, \dots, n.$$

4. The One-Dimensional Case

Theorem 1. *Let* $\Omega = (a, b)$ *be a bounded open interval in* $\mathbb{R}$, $p \in (1, \infty)$ *and suppose that the assumptions of Propositions 2 and 3 hold. Suppose moreover that the data* $z \in Y$ *are given such that, for every measurable set* $E \subset (a, b)$ *with* $\operatorname{meas}(E) > 0$*, the equation*

$$K^*J(Ku - z) = 0 \text{ on } E \tag{4.1}$$

does not have a solution $u \in BV(a, b)$. *Let* $\overline{u}$ *denote the solution to* (P). *Then we have*

$$(\overline{u}')_a = 0.$$

Hence $\overline{u}' = (\overline{u}')_s \in \mathcal{P}$*, and* $\overline{u}' = 0$ *almost everywhere on* (a, b).

Proof. We use the optimality condition (O_1'). In the one-dimensional case, we have $\hat{\lambda} \in W^{1,q}(a, b)$ and

$$\alpha\,\hat{\lambda}' = K^*J(K\overline{u} - z) \in L^q(a, b). \tag{4.2}$$

Suppose that the absolute continuous part $(\overline{u}')_a$ in the Lebesgue decomposition $\overline{u}' = (\overline{u}')_a + (\overline{u}')_s$ of $\overline{u}'$ does not vanish. Let us assume without loss of generality that $(\overline{u}')_a > 0$ on some measurable set E with $\operatorname{meas}(E) > 0$. Then, by ($O_2'$), we have $\hat{\lambda}(x) = 1$ on E and hence $\hat{\lambda}'(x) = 0$ almost everywhere on E (cf. Gilbarg and Trudinger [8, Lem.7.7,p.145]). If we insert this into (4.2), we see that $\overline{u}$ solves (4.1) on some set of positive Lebesgue measure in contradiction to our assumption. Hence $(\overline{u}')_a = 0$ and thus $\overline{u}' = (\overline{u}')_s \in \mathcal{P}$. □

Remark 5. Theorem 1 states that, under the certain conditions on the data z, the measure $\overline{u}'$ is concentrated on a subset of (a, b), which is of Lebesgue measure 0. In this sense, we can say that $\overline{u}' = 0$ almost everywhere on (a, b), and thus " $\overline{u}$ is constant almost everywhere".

Remark 6. It is a common situation for inverse problems that the data z are given by some noisy measurement of 'ideal' data $z_0 \in K(BV(a, b))$. Typically $K(BV(a, b))$ is dense in Y but does not coincide with Y and the perturbed data z are outside this range. Therefore the optimization problem

$$\text{minimize} \|Ku - z\|^2 \text{ over } u \in BV(a, b) \tag{4.3}$$

does not have a solution, and consequently, the optimality condition

$$K^* J(Ku - z) = 0 \tag{4.4}$$

does not have a solution $u \in BV(a,b)$. Condition (4.1) can therefore be seen as a generalization of $z \notin K(BV(a,b))$.

We investigate more closely two special situations which are important in practical applications: the case $K = \mathrm{id}$, which corresponds to the problem of noise removal from a noise-corrupted signal, and the case where K is some smoothing integral operator, which corresponds to the problem of deblurring a blurred signal.

Proposition 5 (Denoising)**.** *Suppose $Y = L^2(a,b)$, $p = q = 2$, and $K = \mathrm{id}$. Assume moreover that $z \in L^2(a,b)$ is given such that for any $E \subset (a,b)$ with positive Lebesgue measure, z does not coincide with some function $u \in BV(a,b)$ on E. Let $\overline{u}$ denote the solution to* (P)*. Then we have $(\overline{u}')_a = 0$, and hence $\overline{u}' = (\overline{u}')_s \in \mathcal{P}$.*

Proof. In this special situation, we have $K = K^* = \mathrm{id}$, and also $J = \mathrm{id}$, (the duality map of the Hilbert space $L^2(a,b)$). Condition (4.1) then reads as: For any $E \subset (a,b)$ of positive measure, $u - z = 0$ on E does not have a solution u $\in BV(a,b)$. This is the condition formulated in the Proposition. □

In the deblurring case, we find that $\overline{u}'$ is a singular measure, under even weaker assumptions on the data z.

Proposition 6 (Deblurring)**.** *Suppose $Y = L^2(a,b)$, $p = q = 2$ and*

$$Ku = \int_a^b k(x - \cdot)\, u(x)\, dx,$$

with k analytic, and $k(x) = k(-x)$. Suppose also that K is injective. Assume moreover that

$$z \notin K(BV(a,b)). \tag{4.5}$$

Let $\overline{u}$ denote the solution to (P)*. Then we have $(\overline{u}')_a = 0$, and hence $\overline{u}' = (\overline{u}')_s \in \mathcal{P}$.*

Proof. It is easy to see that, under the above assumptions K is selfadjoint on $L^2(a,b)$, and, as in Proposition 5, $J = \mathrm{id}$. Condition (4.1) then reads as: For any $E \subset (a,b)$ of positive measure,

$$K(Ku - z) = 0 \text{ on } E \text{ does not have a solution u } \in BV(a,b).$$

Suppose this condition was not satisfied, i.e. there exists $u \in BV(a,b)$ such that $K(Ku - z) = 0$ on E for some set E of positive measure. Since k is analytic, we find that $K(Ku-z)$ is an analytic function, which is 0 on some set of positive measure. This, however, means $K(Ku-z) \equiv 0$ on (a,b), and, by the injectivity of K, we obtain $Ku - z \equiv 0$ on (a,b), i.e. $z \in K(BV(a,b))$, in contradiction to our assumption. □

5. The Two-Dimensional Case

In dimension one, an important point in the proof of Theorem 1 is the fact that we have $\hat{\lambda}' \in L^q$, hence $\hat{\lambda} \in W^{1,q}$, and we can use Lemma 7.7 in [8] to conclude that the distributional derivative of $\hat{\lambda}$ is zero almost everywhere on the set where $\hat{\lambda}$ is constant. In dimension two or higher, we only have $\operatorname{div} \hat{\lambda} \in L^q(\Omega)$, (cf. (O_1')) which is weaker than $\hat{\lambda} \in W_2^{1,1}(\Omega)$. We therefore cannot use the same argument, and a closer inspection shows that the reasoning in [8] cannot easily be adopted to the present situation. In the two-dimensional situation we therefore get only a weaker result, claiming that the gradient $\nabla \overline{u}$ of the solution cannot have two non-zero components on any *open* subset of Ω. This corresponds to the observation in the introductory section that, at each point, the gradient of the solution is parallel to one of the coordinate directions.

Theorem 2. *Suppose $\Omega \subset \mathbb{R}^2$ is an open domain, $p \in (1, \infty)$ and suppose that the assumptions of Propositions 2 and 3 hold. Suppose moreover that the data $z \in Y$ are given such that, for every* open *set $U \subset \Omega$, the equation*

$$K^* J(Ku - z) = 0 \text{ on } U \tag{5.1}$$

does not have a solution $u \in BV(\Omega)$. *Let $\overline{u}$ denote the solution to* (P). *Then, there is* no open subset Ω' *of Ω on which both components $\left(\frac{\partial \overline{u}}{\partial x_i}\right)_a$, $i = 1, 2$ have constant, non-zero sign.*

Proof. We use the optimality condition (O_1'). Suppose, we can find an open $\Omega' \subset \Omega$, where $\left(\frac{\partial \overline{u}}{\partial x_i}\right)_a$ has constant sign for $i = 1, 2$. Suppose, without loss of generality, that both components are positive. Then we have, by (O_2'), $\hat{\lambda}_i = 1$ on Ω' and therefore $\operatorname{div} \hat{\lambda} = 0$ on Ω'. Thus, we have $K^* J(Ku - z) = 0$ on Ω', contradicting our assumption. □

Remark 7. The following conjecture is the analogue to Lemma 7.7 in [8] for vectorfields with divergence in $L^q(\Omega)$. If the conjecture holds, we are able to prove a stronger result.

Conjecture 1. *Let $\lambda \in L_2^\infty(\Omega)$, with* $\operatorname{div} \lambda \in L^q(\Omega)$ *for some $1 < q < \infty$. Suppose that $\lambda =$* const *on some measurable set $E \subset \Omega$. Then* $\operatorname{div} \lambda = 0$ *almost everywhere on E.*

We consider again the solution $\overline{u}$ of problem (P). If we suppose that the data $z \in Y$ are given such that, for every measurable set $E \subset \Omega$ with $\operatorname{meas}(E) > 0$, the equation

$$K^* J(Ku - z) = 0 \text{ on } E \tag{5.2}$$

does not have a solution $u \in BV(\Omega)$, and under the hypothesis, that Conjecture 1 holds, we get the following result: Almost everywhere on Ω, we have

$$\left(\frac{\partial \overline{u}}{\partial x_1}\right)_a = 0 \text{ or } \left(\frac{\partial \overline{u}}{\partial x_2}\right)_a = 0.$$

Remark 8. The result in Theorem 2 depends on the fact that $\hat{\lambda}_i \in \operatorname{sign}\left(\frac{\partial \overline{u}}{\partial x_i}\right)_a$, which is only true for the anisotropic norm on $BV(\Omega)$, which we have chosen in (2.4). For any other choice of the norm this structural result does not hold in this form. This is different from the situation in dimension one, where we have a unique vector norm.

As in dimension one, we consider the denoising and deblurring problem in more detail. The proofs are completely analogous to the proofs of Propositions 5 and 6 and therefore omitted.

Proposition 7 (Denoising)**.** *Suppose $Y = L^2(\Omega)$, $p = q = 2$, and $K = \mathrm{id}$. Assume moreover that $z \in L^2(\Omega)$ is given such that for any open subset $U \subset \Omega$, z does not coincide with some function $u \in BV(\Omega)$ on U. Let $\overline{u}$ denote the solution of problem* (P)*. Then, there is* no open subset *Ω' of Ω on which both components $\left(\frac{\partial \overline{u}}{\partial x_i}\right)_a$, $i = 1, 2$ have constant, non-zero sign.*

Proposition 8 (Deblurring)**.** *Suppose $Y = L^2(a, b)$, $p = q = 2$ and*

$$Ku = \int_\Omega k(x - \cdot)\, u(x)\, dx,$$

with k analytic, and $k(x) = k(-x)$. Suppose also that K is injective. Assume moreover that

$$z \notin K(BV(\Omega)). \tag{5.3}$$

Let $\overline{u}$ denote the solution of problem (P)*. Then, there is* no open subset *Ω' of Ω on which both components $\left(\frac{\partial \overline{u}}{\partial x_i}\right)_a$, $i = 1, 2$ have constant, non-zero sign.*

Remark 9. It is obvious how the two-dimensional results can be generalized to $n > 2$ dimensions. As for $n = 2$, we cannot have that all components of the gradient are not zero on any open subset of Ω.

References

[1] R. Acar and C. R. Vogel, *Analysis of bounded variation penalty methods for ill-posed problems*, Inverse Problems, 10 (1994), pp. 1217–1229.

[2] V. Barbu, *Analysis and Control of Nonlinear Infinite Dimensional Systems*, vol. 190 of Mathematics in Science and Engineering, Academic Press, Boston, 1993.

[3] A. Chambolle and P. L. Lions, *Image recovery via total variation minimization and related problems*, Numerische Mathematik, 76 (1997), pp. 167–188.

[4] G. Chavent and K. Kunisch, *Regularization of linear least squares problems by total bounded variation*, ESIAM, Control, Optimisation and Calculus of Variation, 2 (1997), pp. 359–376.

[5] D. Dobson and O. Scherzer, *Analysis of regularized total variation penalty methods for denoising*, Inverse Problems, 12 (1996), pp. 601–617.

[6] D. C. Dobson and F. Santosa, *Recovery of blocky images from noisy and blurred data*, SIAM J. Appl. Math., 56 (1996), pp. 1181–1192.

[7] I. Ekeland and T. Turnbull, *Infinite-Dimensional Optimization and Convexity*, Chicago Lectures in Mathematics, The University of Chicago Press, Chicago and London, 1983.

[8] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, vol. 224 of Grundlehren der mathematischen Wissenschaften, Springer Verlag, Berlin, 1977.

[9] E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, vol. 80 of Monographs in Mathematics, Birkhäuser, Boston, 1984.

[10] K. Ito and K. Kunisch, *An active set strategy based on the augmented lagrantian formulation for image restauration*, RAIRO, Math. Mod. and Numer. Anal., 33 (1999), pp. 1–21.

[11] K. Ito and K. Kunisch, *Restoration of edge - flat - grey scale images*, Preprint, (1999).

[12] M. Z. Nashed and O. Scherzer, *Least squares and bounded variation regularization with nondifferentiable functionals*, Numer. Funct. Anal and Optimiz., 19 (1998), pp. 873–901.

[13] L. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithm*, Physica D, 60 (1992), pp. 259–268.

[14] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York – St Louis – San Francisco, third ed., 1987.

[15] C. Vogel and M. Oman, *Iterative methods for total variation denoising*, SIAM J. Sci. Comp., 17 (1996), pp. 227–238.

[16] K. Yosida, *Functional Analysis*, vol. 123 of Grundlehren der mathematischen Wissenschaften, Springer-Verlag, Berlin–Heidelberg–New York, fourth ed., 1974.

[17] W. P. Ziemer, *Weakly Differentiable Functions*, vol. 120 of Graduate Texts in Mathematics, Springer Verlag, New York, 1989.